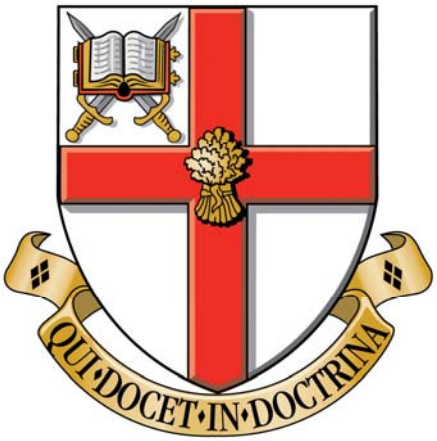

University of Chester

NUMERICAL SOLUTION OF MULTI-ORDER FRACTIONAL DIFFERENTIAL EQUATIONS

KAI DIETHELM* AND NEVILLE J. FORD†

Abstract. We consider the numerical solution of (possibly nonlinear) fractional differential equations of the form

$$y^{(\alpha)}(t) = f(t, y(t), y^{(\beta_1)}(t), y^{(\beta_2)}(t), \ldots, y^{(\beta_n)}(t))$$

with $\alpha > \beta_1 > \beta_2 > \ldots > \beta_n$ and $\alpha - \beta_1 \leq 1$, $\beta_j - \beta_{j+1} \leq 1$, $0 < \beta_n \leq 1$, and where the derivatives are understood in the Caputo sense.

We begin by discussing the questions of existence and uniqueness of solutions, and we investigate how the solutions depend on the given data. We propose convergent and stable numerical methods based on a nearly equivalent system of fractional differential equations of order not exceeding β_n. We give particular emphasis to the practically important linear case and also present results for nonlinear problems.

Key words. Multi-term fractional differential equation, Caputo derivative, existence, uniqueness, structural stability, Adams method

AMS subject classifications. Primary 65L05; secondary 65L06, 65R20, 26A33, 34A12, 34A45, 34D30.

1. Introduction. Numerical methods for the solution of linear fractional differential equations involving only one fractional derivative are well established (see for example [1, 2, 5, 16]). The work [3] contains many references to fractional integration methods and related theory although it does not explicitly address fractional differential equations. There have been some attempts to solve *linear* problems with multiple fractional derivatives (the so-called multi-term equations) [7, 12, 19, 21] but a complete analysis has not been given so far. Nonlinear equations have received rather less attention in the literature, partly because many of the model equations proposed have been linear. Indeed, some writers have proposed that the use of fractional differential equations in a model can avoid altogether the need to introduce nonlinearity. More recently, applications have included classes of nonlinear fractional differential equations (see, for example [9]) and this motivates us to consider their effective numerical solution. To our knowledge this paper presents the first viable numerical method for the solution of nonlinear multi-order fractional differential equations.

We are concerned with providing good quality algorithms for the solution of multi-order fractional differential equations of the general (possibly nonlinear) form

$$y^{(\alpha)}(t) = f(t, y(t), y^{(\beta_1)}(t), y^{(\beta_2)}(t), \ldots, y^{(\beta_n)}(t)) \tag{1.1}$$

where $\alpha > \beta_1 > \beta_2 > \ldots > \beta_n$ and $\alpha - \beta_1 \leq 1$, $\beta_j - \beta_{j+1} \leq 1$, $0 < \beta_n \leq 1$, and its linear special case

$$y^{(\alpha)}(t) = \lambda_0 y(t) + \sum_{j=1}^{n} \lambda_j y^{(\beta_j)}(t) + f(t). \tag{1.2}$$

*Institut für Angewandte Mathematik, Technische Universität Braunschweig, Pockelsstraße 14, 38106 Braunschweig, Germany (K.Diethelm@tu-bs.de)

†Department of Mathematics, Chester College, Parkgate Road, Chester CH1 4BJ, United Kingdom (njford@chester.ac.uk). Corresponding author. Neville Ford is a member of the Manchester Centre for Computational Mathematics

Our approach is based on a generalisation of an approach commonly employed in the solution of ordinary differential equations of order two or above, where the equation is converted to a system of equations of order one. In the case of fractional order equations, the analysis has some unique features. We consider the linear equation first and then develop our theory for nonlinear problems. One advantage of treating the linear equation separately is that there is an explicit expression for the solution and this provides useful additional insight.

In Section 2 we introduce the notation and definitions and discuss the form of initial conditions we shall use. This is particularly important with fractional derivatives because there are several definitions available and they have some fundamental differences. Indeed, it turns out that the form of initial conditions chosen is a key element in the wide applicability of the approach we have developed. We also review and generalize our recent work (see [7]) in which we proposed methods for the solution of the (multi-term) Bagley-Torvik equation based on the conversion of the problem to a system.

In Section 3, we analyse the linear equation. We show that *any* multi-term linear equation may be approximated arbitrarily closely by a system of linear fractional differential equations of a single order. This allows us to invoke the powerful techniques from Section 2 to obtain an arbitrarily good approximation for the unknown solution.

In Section 4, we turn to the nonlinear problem. After a brief discussion of some basic theoretical results, we show that the conversion of a nonlinear multi-order equation to an approximating system of single-order equations is possible in this case too.

In Section 5 we present our results based on the use of an algorithm for solving the resulting nonlinear system.

2. Basic ideas and definitions. We recall the definition of the Riemann-Liouville differential operators of fractional order $q > 0$,

$$D^q y(t) := \frac{1}{\Gamma(m-q)} \frac{d^m}{dt^m} \int_0^t \frac{y(u)}{(t-u)^{q-m+1}} du$$

where m is the integer defined by $m-1 < q < m$ (see [20, 22]) which arise in the modelling of many physical phenomena [9, 14, 15, 17, 18].

The standard approach [22, §42], is to define the initial conditions corresponding to (1.1) or (1.2) in the form

$$\frac{d^{q-k}}{dt^{q-k}} y(t)|_{t=0+} = b_k, \qquad k = 1, 2, \ldots, m = \lfloor q+1 \rfloor,$$

with given values b_k. Thus we are forced to specify some fractional derivatives of the function y. In practical applications, these values are frequently not available, and it may not even be clear what their physical meaning is (see [9]). Therefore Caputo [4] has suggested that one should incorporate the classical derivatives (of integer order) of the function y, as they are commonly used in initial value problems with integer-order equations, into the fractional-order equation, giving

$$D_*^q y(t) := D^q (y - T_{m-1}[y])(t) = f(t, y(t)), \tag{2.1a}$$

where $T_{m-1}[y]$ is the Taylor polynomial of order $(m-1)$ for y, centered at 0. For $q \in \mathbb{N}$, one simply defines D_*^q to be the usual differential operator of order q. Then, one can specify the initial conditions in the classical form

$$y^{(k)}(0) = y_0^{(k)}, \qquad k = 0, 1, \ldots, m-1. \tag{2.1b}$$

It is derivatives of the Caputo type defined in (2.1a) and initial conditions of the form (2.1b) that we shall apply in this paper. As we remarked in [7], the use of Caputo initial conditions allows a more general application of our technique than would be the case for the Riemann-Liouville formulation. The two formulations coincide when the initial conditions are zero. Some authors (see, for example, [21]) remark that zero initial conditions are the most natural case to arise in applications and are also the special case where all available definitions of the fractional derivative coincide.

As we saw in our recent paper [7] it can be a simple matter to convert some linear equations with commensurate multiple fractional derivatives into a linear system of fractional differential equations of low order. One can then show that the application of (for example) fractional linear multistep methods to the solution of the resulting system is exactly equivalent to the solution of the problem in its original form by the corresponding fractional linear multistep methods. One can also show that this scheme for the linear system inherits the properties (derived by Lubich [16]) for the scalar case and that there results an effective algorithm for the solution of the equation.

To make our approach clear consider an example. For the equation

$$D_*^{\alpha} y(t) = f(t, y(t), D_*^{\beta_1} y(t), D_*^{\beta_2} y(t), \dots, D_*^{\beta_n} y(t)) \tag{2.2}$$

subject to the initial conditions

$$y^{(k)}(0) = y_0^{(k)}, \qquad k = 0, 1, \dots, \lceil \alpha \rceil - 1 \tag{2.3}$$

we assume $\alpha > \beta_1 > \beta_2 > \ldots > \beta_n$ and $\alpha - \beta_1 \leq 1$, $\beta_j - \beta_{j+1} \leq 1$, $0 < \beta_n \leq 1$ and $\alpha, \beta_j \in \mathbb{Q}$. Let M be the least common multiple of the denominators of α, β_1, $\dots$, β_n, and set $\gamma := 1/M$ and $N := M\alpha$. We have the following theorem on equivalence of a nonlinear system:

THEOREM 2.1. *The equation* (2.2)*, equipped with the initial conditions* (2.3)*, is equivalent to the system of equations*

$$\begin{aligned} D_*^{\gamma} y_1(t) &= y_2(t), \\ D_*^{\gamma} y_2(t) &= y_3(t), \\ D_*^{\gamma} y_3(t) &= y_4(t), \\ &\vdots \\ D_*^{\gamma} y_{N-1}(t) &= y_N(t), \\ D_*^{\gamma} y_N(t) &= f(t, y_{\bar{\beta}_1/\gamma+1}(t), \dots, y_{\bar{\beta}_n/\gamma+1}(t)), \end{aligned} \tag{2.4}$$

together with the initial conditions

$$y_j(0) = \begin{cases} y_0^{(k)} & \text{if } j = kM + 1 \text{ with some } k \in \mathbb{N}, \\ 0 & \text{else}, \end{cases} \tag{2.5}$$

in the following sense.

1. *Whenever* $Y := (y_1, \dots, y_N)^T$ *with* $y_1 \in C^{\lceil \alpha \rceil}[0, b]$ *for some* $b > 0$ *is the solution of the system* (2.4)*, equipped with the corresponding initial conditions, the function* $y := y_1$ *solves the multi-term equation* (2.2)*, and it satisfies the initial conditions* (2.3)*.*
2. *Whenever* $y \in C^{\lceil \alpha \rceil}[0, b]$ *is a solution of the multi-term equation* (2.2) *satisfying the initial conditions* (2.3)*, the vector-valued function* $Y := (y_1, \dots y_N)^T := (y, D_*^{\gamma} y, D_*^{2\gamma} y, \dots, D_*^{(N-1)\gamma} y)^T$ *satisfies the system* (2.4) *and the initial conditions* (2.5)*.*

This statement is a simple generalisation of [7, Thm. 2.1]. The proof can be carried over easily.

We note in particular that the system (2.4) is linear if the given equation (2.2) is linear.

3. The linear theory. Theorem 2.1 has shown that the conversion of a fractional differential equation with multiple fractional derivatives to a system follows the corresponding idea for integer order equations. However for fractional order equations one requires the assumption of commensuracy of the orders. It has been remarked elsewhere that when the orders are not commensurate then there is no system of fractional equations that *exactly* corresponds to the original problem. To overcome this difficulty we use the well known fact that any real number can be approximated arbitrarily closely by a rational number and therefore one can *approximate* any linear fractional differential equation with multiple fractional derivatives by an equation whose orders are as close as we choose to the original orders, and yet whose orders are commensurate (in fact rational) — a property that will apply in any case as soon as the orders are stored in a computer. In this section we prove that the problem is *structurally stable* in the sense that the solution to the approximating linear system that results lies close to the solution of the original linear fractional differential equation. To be precise, we will prove that the error in the solution that results from introducing these perturbed orders of derivative is of the same order of magnitude as the perturbation introduced (cf. [6]).

REMARK 3.1. *Without loss of generality, we assume that every non rational order is approximated by a nearby rational order in the same interval* $(\nu, \nu+1), \nu \in \mathbb{N}$. *If we did not make this assumption, we could need different initial conditions in the perturbed equations compared with the original problem.*

Our tool in this part of the analysis is the fractional Green's function (see, for example, [21]). This allows us to give an explicit representation of the solution to both the original problem and to the approximation and leads to our conclusion. The analytical solution of the original problem takes the form:

$$y(t) = \int_0^t G(t-u) f(u) du$$

with

$$\begin{aligned} G(t) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} & \sum_{k_0+k_1+\ldots+k_{n-2}=m, k_i \geq 0} (m; k_0, \ldots, k_{n-2}) \\ & \times \prod_{i=0}^{n-2} (-\lambda_{n-i})^{k_i} t^{(\alpha-\beta_1)m+\alpha+\sum_{j=0}^{n-2}(\beta_1-\beta_{n-j})k_j-1} \\ & \times E^{(m)}_{\alpha-\beta_1, \alpha+\sum_{j=0}^{n-2}(\beta_1-\beta_{n-j})k_j} \left(\lambda_1 t^{\alpha-\beta_1}\right) \end{aligned}$$

where $(m; k_0, \ldots, k_{n-2}) = m! / \prod_{i=0}^{n-2}(k_i!)$ is the multinomial coefficient and $E^{(k)}_{\lambda,\mu}$ is the kth derivative of the Mittag-Leffler function with parameters λ and μ, given by

$$E^{(k)}_{\lambda,\mu}(t) = \sum_{j=0}^{\infty} \frac{(j+k)! t^j}{j! \Gamma(\lambda j + \lambda k + \mu)}.$$

For the perturbed problem, the solution takes the form

$$\tilde{y}(t) = \int_0^t \tilde{G}(t-u)f(u)du$$

with

$$\begin{aligned}\tilde{G}(t) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} \sum_{k_0+k_1+\ldots+k_{n-2}=m, k_i\geq 0} (m; k_0, \ldots, k_{n-2}) \\ \times \prod_{i=0}^{n-2} (-\lambda_{n-i})^{k_i} t^{(\tilde{\alpha}-\tilde{\beta}_1)m+\tilde{\alpha}+\sum_{j=0}^{n-2}(\tilde{\beta}_1-\tilde{\beta}n-j)k_j-1} \\ \times E^{(m)}_{\tilde{\alpha}-\tilde{\beta}_1, \tilde{\alpha}+\sum_{j=0}^{n-2}(\tilde{\beta}_1-\tilde{\beta}_{n-j})k_j} \left(\lambda_1 t^{\tilde{\alpha}-\tilde{\beta}_1}\right).\end{aligned}$$

It follows that the perturbation in the solution $z = y - \tilde{y}$ satisfies

$$z(t) = \int_0^t \hat{G}(t-u)f(u)du$$

with $\hat{G}(t) = G(t) - \tilde{G}(t)$. We can conclude that, over any finite time interval $[0, T]$, the value z is bounded by $\|\hat{G}\|_\infty \cdot \|f\|_\infty T$, and using a Gronwall-type argument it is straightforward to show that $\|\hat{G}\|_\infty = \mathcal{O}(\epsilon)$ where

$$\epsilon = \max\{\alpha - \tilde{\alpha}, \beta_1 - \tilde{\beta}_1, \ldots, \beta_n - \tilde{\beta}_n\}.$$

We summarise this in the following Theorem:

THEOREM 3.1. *Let y be the solution of*

$$D^\alpha_* y(t) = \lambda_0 y(t) + \sum_{j=1}^{n} \lambda_j D^{\beta_j}_* y(t) + f(t)$$

with initial conditions $y^{(k)}(0) = y_0^{(k)}$, $k = 0, 1, \ldots, \lceil\alpha\rceil - 1$, and let z be the solution of

$$D^{\tilde{\alpha}}_* z(t) = \lambda_0 y(t) + \sum_{j=1}^{n} \lambda_j D^{\tilde{\beta}_j}_* z(t) + f(t)$$

with initial conditions $z^{(k)}(0) = y_0^{(k)}$, $k = 0, 1, \ldots, \lceil\alpha\rceil - 1$ where $|\alpha - \tilde{\alpha}| < \epsilon, |\beta_j - \tilde{\beta}_j| < \epsilon$. For $T < \infty$, we have

$$\|y - z\|_{L_\infty[0,T]} = O(\epsilon), \quad \epsilon \to 0.$$

4. The nonlinear problem. We adopt the same general approach for the nonlinear problem. Here we begin with a proof that the nonlinear equation

$$D^\alpha_* y(t) = f(t, y(t), D^{\beta_1}_* y(t), D^{\beta_2}_* y(t), \ldots, D^{\beta_n}_* y(t)) \tag{4.1}$$

subject to the initial conditions

$$y^{(k)}(0) = y_0^{(k)}, \qquad k = 0, 1, \ldots, \lceil\alpha\rceil - 1 \tag{4.2}$$

has (under natural Lipschitz conditions imposed on f) a unique continuous solution. Note that this initial value problem coincides with the problem (2.2) and (2.3) that we considered in Theorem 2.1, but with the difference that for the existence and uniqueness proof below we do not need the number-theoretic hypotheses assumed in that theorem. Those properties are only important when discussing the conversion of the multi-order equation to a system of single-order equations. We recall that our previous paper [6] gave the following existence and uniqueness results for the equation

$$D_*^\alpha y(t) = f(t, y(t)).$$

LEMMA 4.1 (Existence (see [6])). *Assume that* $\mathcal{D} := [0, \chi^*] \times [y_0^{(0)} - a, y_0^{(0)} + a]$ *with some* $\chi^* > 0$ *and some* $a > 0$, *and let the function* $f : \mathcal{D} \to \mathbb{R}$ *be continuous. Furthermore, define* $\chi := \min\{\chi^*, (a\Gamma(q+1)/\|f\|_\infty)^{1/q}\}$. *Then, there exists a function* $y : [0, \chi] \to \mathbb{R}$ *solving the initial value problem* (2.1).

LEMMA 4.2 (Uniqueness (see [6])). *Assume that* $\mathcal{D} := [0, \chi^*] \times [y_0^{(0)} - a, y_0^{(0)} + a]$ *with some* $\chi^* > 0$ *and some* $a > 0$. *Furthermore, let the function* $f : \mathcal{D} \to \mathbb{R}$ *be bounded on* $\mathcal{D}$ *and fulfil a Lipschitz condition with respect to the second variable, i.e.*

$$|f(x, y) - f(x, z)| \le L|y - z|$$

with some constant $L > 0$ *independent of* x, y, *and* z. *Then, denoting* χ *as in Lemma* 4.1, *there exists at most one function* $y : [0, \chi] \to \mathbb{R}$ *solving the initial value problem* (2.1).

The generalisation of Lemmas 4.1 and 4.2 to vector-valued functions y is immediate.

For the nonlinear equation with multiple fractional derivatives of commensurate order, we begin by converting (as in Section 2) to a system of nonlinear fractional differential equations of low order. As above, M is the least common multiple of the denominators of α, β_1, ..., β_n, and we set $\gamma := 1/M$ and $N := M\alpha$.

Lemmas 4.1 and 4.2 then apply to the system (2.4) with initial conditions (2.5) yielding an existence-uniqueness theorem for the commensurate multi-order problem:

THEOREM 4.3 (Existence and uniqueness (multi-term, commensurate orders)). *Let the continuous function* f *in* (4.1) *satisfy a uniform Lipschitz condition with Lipschitz constant* L *in all its arguments except for the first on a suitable domain* $\mathcal{D}$. *Assume further that the orders* α, β_j *are rational. Then the equation* (4.1) *subject to* (4.2) *has a unique continuous solution on an interval* $[0, T]$ *of the real line.*

Next we derive a Gronwall-type result (see, for example, [3] for a discussion of similar results that apply to integral equations). We show that, under small variations in the orders β_j in (4.1), we can give a uniform bound on the change in the solution on any closed bounded interval $[0, T]$. We state and prove the result for an equation with two terms, but the generalisation to multi-term equations is straightforward.

THEOREM 4.4 (Gronwall-type result for a two-term equation). *Let* $\alpha > \beta, \tilde{\beta} > 0$ *be chosen so that the equations*

$$D_*^\alpha y(t) = f(t, y(t), D_*^\beta y(t)) \tag{4.3}$$

subject to the initial conditions

$$y(0) = y_0, y'(0) = y_0', \ldots, y^{(\lceil\alpha\rceil-1)}(0) = y_0^{(\lceil\alpha\rceil-1)} \tag{4.4}$$

and

$$D_*^\alpha z(t) = f(t, z(t), D_*^{\tilde{\beta}} z(t)) \tag{4.5}$$

subject to the same initial conditions

$$z(0) = y_0, z'(0) = y_0', \dots, z^{(\lceil\alpha\rceil-1)}(0) = y_0^{(\lceil\alpha\rceil-1)} \tag{4.6}$$

(where f satisfies a Lipschitz condition in its second and third arguments on a suitable domain) have unique continuous solutions y, z. We assume further that $\lfloor\beta\rfloor = \lfloor\tilde{\beta}\rfloor$. Then $Y = y - z$ satisfies a relation of the form

$$Y(t) \leq K|\beta - \tilde{\beta}| + \int_0^t \Lambda Y(s)ds$$

and it follows that there exist constants K and Λ such that

$$|y(t) - z(t)| \leq K|\beta - \tilde{\beta}| \exp(\Lambda T)$$

for all $t \in [0, T]$.

REMARK 4.1. *Note that if $\alpha, \beta, \tilde{\beta}$ are all rational then the equations (4.3), (4.5) each have a unique continuous solution and the conclusions of Theorem 2.1 apply.*

Proof. We write the solutions y and z in the form of the equivalent Volterra integral equations:

$$y(t) = \sum_{j=0}^{\lceil\alpha\rceil-1} \frac{y_j}{j!} t^j + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(s, y(s), D_*^{\beta} y(s))ds \tag{4.7}$$

and

$$z(t) = \sum_{j=0}^{\lceil\alpha\rceil-1} \frac{y_j}{j!} t^j + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(s, z(s), D_*^{\tilde{\beta}} z(s))ds \tag{4.8}$$

and we fix $T > 0$. Subtracting we obtain the relation

$$\begin{aligned} y(t) - z(t) = &\frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \left(f(s, y(s), D_*^{\beta} y(s)) - f(s, y(s), D_*^{\tilde{\beta}} y(s)) \right) ds \\ &+ \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \left(f(s, y(s), D_*^{\tilde{\beta}} y(s)) - f(s, z(s), D_*^{\tilde{\beta}} z(s)) \right) ds. \end{aligned}$$

Now, with $m \in \mathbb{N}$ chosen so that $m-1 < \beta, \tilde{\beta} < m$, and bearing in mind that y is the unique solution to (4.3) on $[0, T]$ we can estimate (using the Lipschitz condition on f and the definition of the Caputo derivative) the first term on the right hand side:

$$\left| \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \left(f(s, y(s), D_*^{\beta} y(s)) - f(s, y(s), D_*^{\tilde{\beta}} y(s)) \right) ds \right| \leq K|\beta - \tilde{\beta}|$$

uniformly for $t \in [0, T]$. Moreover we can use the Lipschitz conditions on f in the second term on the right hand side to give

$$\begin{aligned} &\left| \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \left(f(s, y(s), D_*^{\tilde{\beta}} y(s)) - f(s, z(s), D_*^{\tilde{\beta}} z(s)) \right) ds \right| \\ &\qquad < \int_0^t \Lambda |y(s) - z(s)| ds \end{aligned}$$

by evaluating the integral representations of the fractional derivatives of y and z. If we put $Y(t) = |y(t) - z(t)|$ it follows that

$$Y(t) \leq K|\beta - \tilde{\beta}| + \int_0^t \Lambda Y(s)ds \tag{4.9}$$

Equation (4.9) now allows us to conclude (see, for example [3, p. 38]) that

$$Y(t) \leq K|\beta - \tilde{\beta}| \exp(\Lambda T) \tag{4.10}$$

uniformly for $t \in [0, T]$ and the proof is complete. □

We use the conclusion of Theorem 4.4 several times. First we derive an existence and uniqueness Theorem for more general multi-term nonlinear equations (with non-commensurate multiple derivatives). As before, we give a proof for the two-term equation; the result can be easily generalised for equations with more terms.

THEOREM 4.5 (General Existence-Uniqueness). *Let the function f satisfy a uniform Lipschitz condition in its second and third arguments and be continuous in its first argument. It follows that the equation*

$$D_*^\alpha y(t) = f(t, y(t), D^\beta y(t)) \tag{4.11}$$

where $\alpha > \beta > 0$, subject to the initial conditions

$$y(0) = y_0, y'(0) = y_0', \ldots, y^{(\lfloor \alpha \rfloor)}(0) = y_0^{(\lfloor \alpha \rfloor)}$$

has a unique continuous solution on any finite interval $[0, T]$.

Proof. Case 1: $\alpha, \beta \in \mathbb{Q}$. We observe that the result is already established when α and β are rational because of Theorem 2.1 and Lemmas 4.1 and 4.2.

Case 2: $\alpha \in \mathbb{Q}$, $\beta \notin \mathbb{Q}$. We construct a sequence (β_j) of rational numbers whose limit is β. Without loss of generality, the sequence lies in the interval $(\lfloor \beta \rfloor, \lfloor \beta + 1 \rfloor)$. Clearly by case 1 the equation (4.11) with β replaced in turn by each β_j has a unique continuous solution $y^{[j]}$ on $[0, T]$ whose Caputo derivative of order α is also continuous. We now use equation (4.9) together with the fact that $\beta_j - \beta \to 0$ as $j \to \infty$ to conclude that the sequence of solutions converges uniformly on $[0, T]$ to a continuous function y with $D_*^\alpha y$ also being continuous. It remains to prove that this function y is the solution of (4.11).

To this end, define $r_j(z) := \|D_*^\alpha z - f(\cdot, z(\cdot), D_*^{\beta_j} z(\cdot))\|_{L_\infty[0,T]}$ for any z such that $D_*^\alpha z$ is continuous. We immediately obtain $r_j(y^{[j]}) = 0$ for all j. Since $y^{[j]} \to y$ uniformly as $j \to \infty$ and r is continuous on the space we consider, we find that

$$r_j(y^{[j]}) - r_j(y) \to 0 \qquad \text{as } j \to \infty.$$

Therefore y is a solution of the given initial value problem.

Because of the Lipschitz condition on f, we can prove the uniqueness of the solution by the usual Picard iteration techniques (cf. the proof of [6, Thm. 2.2]).

Case 3: $\alpha \notin \mathbb{Q}$. In this case we use a sequence (α_j) of rational numbers satisfying $\alpha < \alpha_j < \lceil \alpha \rceil$ for all j and $\lim_{j\to\infty} \alpha_j = \alpha$. For each j we can use either case 1 or case 2 (depending on whether β is rational or not) to provide a unique solution to the perturbed equation obtained by replacing α by α_j. Then we proceed as in case 2 to show that the corresponding sequence of solutions converges to the unique solution of the original problem. □

We can now use the conclusion of Theorem 4.4 to derive a theorem on the structural stability of the equation (4.3) even under small perturbations in the orders of derivatives.

THEOREM 4.6 (Structural Stability). *Let y be the solution of*

$$D_*^{\alpha} y(t) = f(t, y(t), D_*^{\beta_1} y(t), D_*^{\beta_2} y(t), \dots, D_*^{\beta_n} y(t))$$

with initial conditions

$$y^{(k)}(0) = y_0^{(k)}, \qquad k = 0, 1, \dots, \lfloor \alpha \rfloor$$

and let z be the solution of

$$D_*^{\tilde{\alpha}} z(t) = f(t, z(t), D_*^{\tilde{\beta}_1} z(t), D_*^{\tilde{\beta}_2} z(t), \dots, D_*^{\tilde{\beta}_n} z(t))$$

with initial conditions

$$z^{(k)}(0) = y_0^{(k)}, \qquad k = 0, 1, \dots, \lfloor \alpha \rfloor$$

where $|\alpha - \tilde{\alpha}| < \epsilon, |\beta_j - \tilde{\beta}_j| < \epsilon$. For $T < \infty$, we have

$$\|y - z\|_{L_\infty[0,T]} = O(\epsilon), \quad \epsilon \to 0$$

The theorem follows from the observation that the difference $y - z$ is a Lipschitz function of $\alpha - \tilde{\alpha}, \beta_1 - \tilde{\beta}_1, \dots, \beta_n - \tilde{\beta}_n$ because of the Gronwall-type Theorem 4.4.

REMARK 4.2. *It follows, by the application of Theorem 4.6, that the solution of any non-commensurate multi-order fractional differential equation may be arbitrarily closely approximated over any finite time interval $[0, T]$ by solutions of equations of rational order (which may in turn be solved by conversion to a system of equations of low order).*

5. A numerical method for general nonlinear equations. Finally we present an Adams-type predictor-corrector formula for the solution of systems of nonlinear equations of low order. Such formulae are well established and are known to give excellent results when applied to classical ordinary differential equations. We further remark that the approach presented here of conversion to a large system of fractional differential equations of low order may be useful in the development of further classes of methods for solving problems of this type.

Our approach follows the outline introduced in the previous sections. Specifically, given a (possibly nonlinear) fractional differential equation of the form

$$D_*^{\tilde{\alpha}} \tilde{y}(t) = f(t, D_*^{\tilde{\beta}_1} \tilde{y}(t), D_*^{\tilde{\beta}_2} \tilde{y}(t), \dots, D_*^{\tilde{\beta}_n} \tilde{y}(t)) \tag{5.1}$$

with initial conditions

$$\tilde{y}^{(k)}(0) = y_0^{(k)}, \qquad k = 0, 1, \dots, \lceil \tilde{\alpha} \rceil - 1,$$

we begin by approximating this equation by

$$D_*^{\alpha} y(t) = f(t, D_*^{\beta_1} y(t), D_*^{\beta_2} y(t), \dots, D_*^{\beta_n}(t)) \tag{5.2}$$

with the same initial conditions, where now the values α, β_1, $\dots$, β_n are rational. We have seen that, by a proper choice of the new parameters, the difference $\|y - \tilde{y}\|_\infty$

of the solutions of these two equations can be made arbitrarily small. Then, in order to solve eq. (5.2) numerically, we first convert the equation into an equivalent system of equations of order γ as in Theorem 2.1. Thus we derive the system (2.4) with initial conditions (2.5). To simplify things, we write this multidimensional initial value problem in the more abstract form

$$D_*^\gamma Y(t) = g(t, Y(t)), \qquad Y(0) = Y_0, \tag{5.3}$$

where $Y : [0,T] \to \mathbb{R}^N$, $Y_0 \in \mathbb{R}^N$, and $g : [0,T] \times \mathbb{R}^N \to \mathbb{R}^N$.

For the solution of this problem, we propose to use the fractional Adams-Bashforth-Moulton scheme introduced in [9] and investigated in a more detailed way in [8]. The scheme is based on rewriting the initial value problem (5.3) as an equivalent fractional integral equation,

$$Y(t) = \sum_{j=0}^{\lceil\alpha\rceil-1} \frac{Y_j}{j!} t^j + \frac{1}{\Gamma(\gamma)} \int_0^t (t-u)^{\gamma-1} g(u, Y(u)) du. \tag{5.4}$$

We then introduce uniformly distributed grid points $t_j := jh$ with $h = T/\nu$ for some $\nu \in \mathbb{N}$ and look at the Volterra equation at these grid points. Here then we find an approximation Y_j, say, for $Y(t_j)$, by a predictor-corrector approach: To obtain the predictor, we replace the integral by a product rectangle quadrature formula (fractional forward Euler method); for the corrector we use the product trapezoidal formula. For full details, including explicit expressions for the weights required in an implementation of the method, we refer to [8, 10, 11].

In our paper [11] we have considered in detail the convergence properties of the method applied here under various alternative assumptions on properties of the solution or the given functions. When $\gamma > 1$ it turns out that, as already stated without proof in [9], we may obtain $O(h^2)$ errors as in the classical case of first-order differential equations. However the construction described in §2 yields that normally we have $0 < \gamma < 1$ since γ is chosen to be as $1/M$ with some integer M. In the case $0 < \gamma < 1$ both our theoretical investigations and our numerical experiments described in [11] suggest that we are unlikely to do better than $O(h^{1+\gamma})$.

As a specific example, we choose the equation

$$D_*^{1.455}\tilde{y}(t) = -t^{0.1} \frac{E_{1.545}(-t)}{E_{1.445}(-t)} \exp(t)\tilde{y}(t) D_*^{0.555}\tilde{y}(t) + \exp(-2t) - [D_*^1 \tilde{y}(t)]^2 \tag{5.5}$$

for $0 \le t \le 1$, equipped with the initial conditions $\tilde{y}(0) = 1$ and $\tilde{y}'(0) = -1$. Here E_α denotes the Mittag-Leffler function with only one parameter α, defined by

$$E_\alpha(z) := \sum_{j=0}^{\infty} \frac{z^j}{\Gamma(\alpha j + 1)}.$$

Direct calculation enabled us to construct this equation in such a way that its exact solution is given by $\tilde{y}(t) = \exp(-t)$. We have chosen an example with differential operators of rational orders because this allows us to compare the approximations obtained by the scheme described above (approximation by other rationals and then application of the Adams method) with the approximations obtained by a direct application of the Adams scheme. This displays the influence of the perturbation of the orders. More details about this feature and examples can be found in [6].

First, we looked at the following approximation of (5.5),

$$D_*^{1.5}y(t) = -t^{0.1}\frac{E_{1.545}(-t)}{E_{1.445}(-t)}\exp(t)y(t)D_*^{0.5}y(t) + \exp(-2t) - [D_*^1 y(t)]^2. \qquad (5.6)$$

Of course, following our strategy the initial conditions that we have to combine with this equation are the same as those originally given for (5.5).

According to Theorem 2.1, this can be converted into an equivalent three-dimensional system (2.4) of order 1/2, where in particular the last equation is

$$D_*^{0.5}y_3(t) = -t^{0.1}\frac{E_{1.545}(-t)}{E_{1.445}(-t)}\exp(t)y_1(t)y_2(t) + \exp(-2t) - [y_3(t)]^2.$$

We have solved this system numerically with the Adams-Bashforth-Moulton algorithm. The resulting errors are displayed in Fig. 5.1. In the calculations, the step sizes were chosen as 1/10, 1/20, and 1/40, respectively.

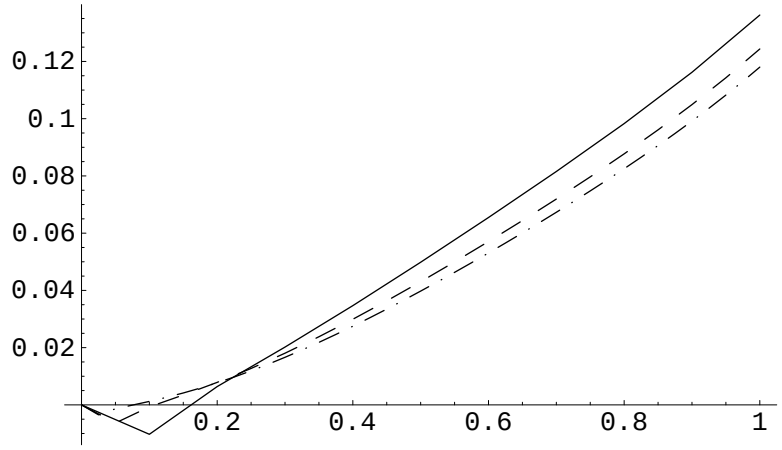

FIG. 5.1. *Approximation errors for the coarse approximation (5.6) (dimension = 3) with* $h = 1/10$ *(continuous line),* $h = 1/20$ *(dashed line),* $h = 1/40$ *(dot-dashed line)*

A look at this figure indicates that the algorithm converges as $h \to 0$. Of course the error does not converge to zero but rather to $\tilde{y} - y$, where $\tilde{y}$ is the (in general unknown) exact solution of (5.5) and y is the exact solution of the differential equation (5.6) that we use as an approximation for (5.5). In view of the fact that the exact solution satisfies $\tilde{y}(1) = \exp(-1) \approx 0.368$, we find a qualitatively correct picture; quantitatively the relative error is about 1/3. A comparison with the results given below for the exact equation (5.5) shows that the method investigated here is extremely simple, especially as far as the requirements of computer memory and run time are concerned. Therefore one may accept this approximation as satisfactory.

Nevertheless, a relative error of 1/3 is too big if one is interested in qualitatively correct data. Therefore we now come to a more precise approximation. To find this approximation, we attempt a better approximation of the original differential equation. To achieve this goal, we perturb the orders of the differential operators by a smaller amount than before. The resulting equation is

$$D_*^{1.45}y(t) = -t^{0.1}\frac{E_{1.545}(-t)}{E_{1.445}(-t)}\exp(t)y(t)D_*^{0.55}y(t) + \exp(-2t) - [D_*^1 y(t)]^2, \qquad (5.7)$$

again equipped with the same initial conditions. Once more we transform this equation to an equivalent system (2.4). This time the order of this system is $\gamma = 1/20$, and the dimension is $N = 29$. We have solved the system with the same numerical method as in the previous case; additionally we have tried the step size $1/80$. In view of the fact that the dimension is now almost ten times as large as in the first case, the computational work involved (i.e. the run time, unless paralellization is used) and the required computer memory also increase by a factor of (approximately) 10. The errors obtained are reported in Fig. 5.2.

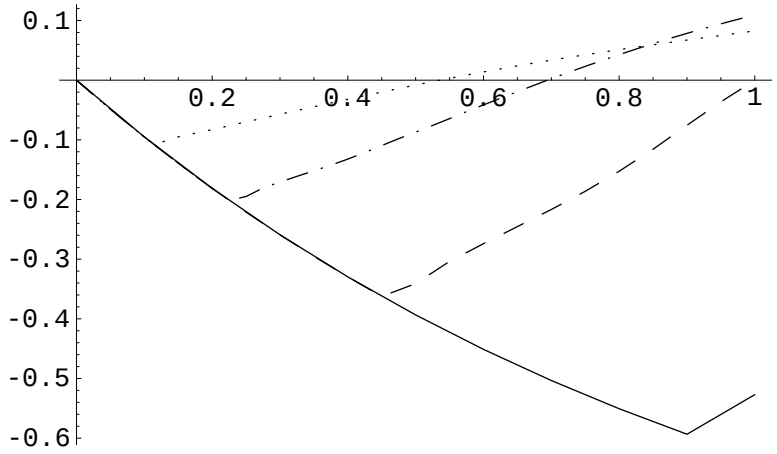

FIG. 5.2. *Approximation errors for the finer approximation (5.7) (dimension = 29) with* $h = 1/10$ *(continuous line),* $h = 1/20$ *(dashed line),* $h = 1/40$ *(dot-dashed line),* $h = 1/80$ *(dotted line)*

At first sight, three observations can be noted: (i) Convergence seems to take place, but much slower than in the previous example; (ii) in the first steps of the algorithm, the error is much larger than in the previous example; (iii) the errors for the various computations (with different step sizes) are identical in the first part of the interval under consideration before they start to behave in different ways. There is a common explanation for these three phenomena. To find this explanation, it is useful to look at the way in which the Adams method interacts with the system (2.4) and with the initial conditions (2.5) for this system. Remember here that the solution that we are interested in is the first component of the vector Y. At the point $t_0 = 0$, we have the exact value (given by the initial condition). Moreover, in view of (2.5), the next $M-1$ components of Y_0 are zero. Now the Adams method determines Y_1 (the approximation for $Y(t_1)$) in the way that first a predictor is calculated. In view of the structure of the right-hand side of the system (2.4), the rth component of the predictor is a linear combination of the rth component of the initial value with the $(r+1)$st components of the $g(t_k, Y_k)$ $(k = 0, 1, \ldots, j-1)$. Specifically, the first component of the predictor is the linear combination of the exact initial value (with coefficient 1) and zero, and hence it is the initial value itself. Moreover, the components of index $2, 3, \ldots, M-1$ of the predictor all have the same form: linear combination of zero and more zeros, and thus they vanish too. The Mth component is the first nonvanishing one. Then we calculate the corrector, and a similar effect happens: The influence of the initial value $Y_M(0)$ is propagated one more row to the $(M-1)$st component of Y_1, but all the previous components of Y_1 coincide with those of Y_0. This pattern continues to work through all the following steps: In every predictor-corrector pair,

the initial value $Y_M(0)$ is propagated two more indices, until at last after $M/2$ steps it reaches the first component. Until this point, the first component remains unchanged, and hence the numerical solution is stuck at the constant value defined by the initial condition. So in the case of our example, the first components of $Y_0, Y_1, \ldots, Y_9$ are all equal to 1, and only when we reach Y_{10}, the numerical solution begins to make progress towards the analytical solution. This behaviour is clearly exhibited in Fig. 5.2. Since this phenomenon is independent of the step size, we find that the adverse effects are reduced with decreasing step sizes.

Of course it must be noted that another reason for the slower convergence is that the order of convergence is $O(h^{1+\gamma})$, and now the order γ of the differential operator is smaller than before. But because of the considerations above we also deduce that the values of h used in our calculations are not sufficiently small to reach the asymptotic stage implicitly contained in this error bound statement.

As a consequence we find that for systems of the form (2.4) it is not useful to use step sizes larger than, say, $T/(4M)$. Thus if we try to improve the accuracy of the approximating differential equation, we are typically forced to increase M, and therefore we must simultaneously decrease the step size. Recalling that, due to the non-local character of the fractional derivatives, the arithmetic complexity of the Adams scheme is $O(Nh^{-2})$ (where N is the dimension of the system that also behaves as $O(M)$), we find the overall result that the complexity of the Adams scheme in its present form depends on M as $O(M^3)$. Some recent results [13] indicate that an adaptive grid may offer a way of decreasing the computational effort.

In order to be able to give a full comparison, we shall now look at the numerical solution of the unperturbed initial value problem (5.5). Since the orders of the differential operators are all rational, our approach is applicable. First we construct the equivalent system (2.4). In this case, the order of the differential operator appearing in this system is $\gamma = 1/200$, and the dimension of the system is $N = 291$. According to our observations concerning the second approximation, there is no point in using the same step sizes as before — this would simply give a constant numerical approximation for the solution. The fact that γ is much smaller than before forces us to decrease the step sizes too. We have therefore chosen h as $1/100$, $1/200$, $1/400$, $1/800$, and $1/1600$ in our experiments. The results are reported in Fig. 5.3.

A look at this figure reveals that convergence does indeed take place. However, we also see that $h = 1/100$ is still far too large to give a reasonable approximation. This confirms the above remark on a sensible choice of h. Actually we find acceptable results from $h = 1/800$ onwards. It is obvious that this leads to much larger requirements concerning computer memory and run time than the crude approximations (5.6) and (5.7).

REFERENCES

[1] C. T. H. Baker and M. S. Derakhshan, *Stability barriers to the construction of $\{\rho, \sigma\}$-reducible and fractional quadrature rules*, in Numerical Integration III, H. Braß and G. Hämmerlin, eds., vol. 85 of Internat. Ser. Numer. Math., Basel, 1988, Birkhäuser, pp. 1–15.

[2] L. Blank, *Numerical treatment of differential equations of fractional order*, Numerical Analysis Report 287, Manchester Centre for Computational Mathematics, 1996.

[3] H. Brunner and P. J. van der Houwen, *The numerical solution of Volterra equations*, North Holland, Amsterdam, 1986.

[4] M. Caputo, *Linear models of dissipation whose Q is almost frequency independent, II*, Geophys. J. Royal Astronom. Soc., 13 (1967), pp. 529–539.

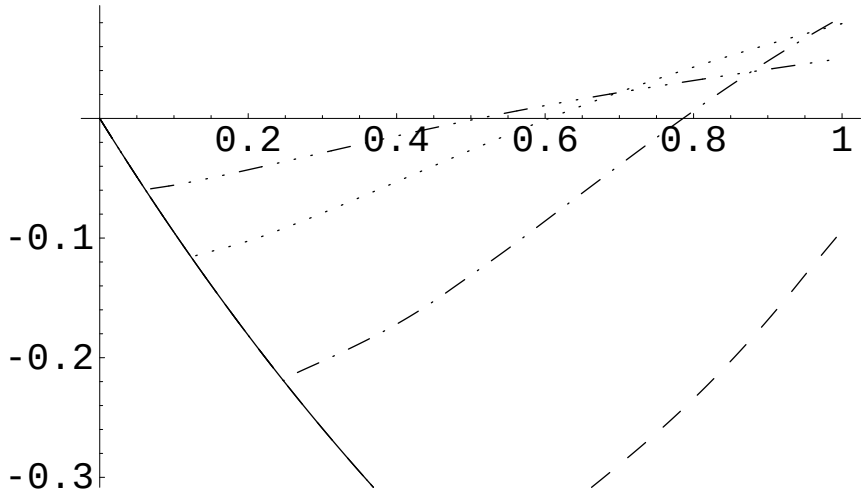

FIG. 5.3. *Approximation errors for the unperturbed system (dimension = 291) with* $h = 1/100$ *(continuous line),* $h = 1/200$ *(dashed line),* $h = 1/400$ *(dot-dashed line),* $h = 1/800$ *(dotted line),* $h = 1/1600$ *(dashed and double-dotted line)*

[5] K. DIETHELM, *An algorithm for the numerical solution of differential equations of fractional order*, Elec. Transact. Numer. Anal., 5 (1997), pp. 1–6.
[6] K. DIETHELM AND N. J. FORD, *Analysis of fractional differential equations*, J. Math. Anal. Appl., 265 (2002), pp. 229–248.
[7] ———, *Numerical solution of the Bagley-Torvik equation*, BIT, to appear (2002).
[8] K. DIETHELM AND A. D. FREED, *The FracPECE subroutine for the numerical solution of differential equations of fractional order*, in Forschung und wissenschaftliches Rechnen 1998, S. Heinzel and T. Plesser, eds., no. 52 in GWDG-Berichte, Göttingen, 1999, Gesellschaft für wissenschaftliche Datenverarbeitung, pp. 57–71.
[9] ———, *On the solution of nonlinear fractional differential equations used in the modeling of viscoplasticity*, in Scientific Computing in Chemical Engineering II — Computational Fluid Dynamics, Reaction Engineering, and Molecular Properties, F. Keil, W. Mackens, H. Voß and J. Werther, eds., Heidelberg, 1999, Springer, pp. 217–224.
[10] K. DIETHELM, A. D. FREED, AND N. J. FORD, *A predictor-corrector approach for the numerical solution of fractional differential equations*, Nonlinear Dynamics, to appear.
[11] ———, *Detailed convergence analysis for a fractional Adams method*, Berichte der Mathematischen Institute der TU Braunschweig, 02/02 (2002).
[12] K. DIETHELM AND Y. LUCHKO, *Numerical solution of linear multi-term differential equations of fractional order*, J. Comput. Anal. Appl., (to appear).
[13] N. J. FORD AND A. C. SIMPSON, *The numerical solution of fractional differential equations: Speed versus accuracy*, Numer. Algorithms, 26 (2001), pp. 333–346.
[14] L. GAUL, P. KLEIN, AND S. KEMPFLE, *Damping description involving fractional operators*, Mech. Systems Signal Processing, 5 (1991), pp. 81–88.
[15] W. G. GLÖCKLE AND T. F. NONNENMACHER, *A fractional calculus approach to self-similar protein dynamics*, Biophysical J., 68 (1995), pp. 46–53.
[16] C. LUBICH, *Fractional linear multistep methods for Abel-Volterra integral equations of the second kind*, Math. Comp., 45 (1985), pp. 463–469.
[17] F. MAINARDI, *Fractional calculus: Some basic problems in continuum and statistical mechanics*, in Fractals and Fractional Calculus in Continuum Mechanics, A. Carpinteri and F. Mainardi, eds., Wien, 1997, Springer, pp. 291–348.
[18] R. METZLER, W. SCHICK, H.-G. KILIAN, AND T. F. NONNENMACHER, *Relaxation in filled polymers: A fractional calculus approach*, J. Chem. Phys., 103 (1995), pp. 7180–7186.
[19] A. R. NKAMNANG, *Diskretisierung von mehrgliedrigen Abelschen Integralgleichungen und gewöhnlichen Differentialgleichungen gebrochener Ordnung*, PhD thesis, Freie Universität, Berlin, 1999. available electronically from http://darwin.inf.fu-berlin.de/1999/23.

[20] K. B. Oldham and J. Spanier, *The Fractional Calculus*, vol. 111 of Mathematics in Science and Engineering, Academic Press, New York, London, 1974.
[21] I. Podlubny, *Fractional Differential Equations*, Academic Press, San Diego, 1999.
[22] S. G. Samko, A. A. Kilbas, and O. I. Marichev, *Fractional Integrals and Derivatives: Theory and Applications*, Gordon and Breach, Yverdon, 1993.